

Comparing orders, rankings, queues, tournaments and lists

Citation for published version (APA):

Can, B., & Storcken, A. J. A. (2015). *Comparing orders, rankings, queues, tournaments and lists*. Maastricht University, Graduate School of Business and Economics. GSBE Research Memoranda No. 020 <https://doi.org/10.26481/umagsb.2015020>

Document status and date:

Published: 01/01/2015

DOI:

[10.26481/umagsb.2015020](https://doi.org/10.26481/umagsb.2015020)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Burak Can, Ton Storcken

**Comparing Orders, Rankings,
Queues, Tournaments and
Lists**

RM/15/020

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O Box 616
NL- 6200 MD Maastricht
The Netherlands

Comparing Orders, Rankings, Queues, Tournaments and Lists

Burak Can* Ton Storcken†

June 18, 2015

Abstract

We study measures on orders that are based on the swaps needed to convert one order into the other. Two classes of such measures are characterized. In one class, the necessary swaps between orders are weighted by their positions in the orders. In the other class, these swaps are weighted by the alternatives involved in these swaps. The results are realized by means of the betweenness condition and several other natural or well-known derivatives. As the axioms are formulated on abstract domains of orders, they apply to all well-known sets of orderings. As a by-product we also show some logical dependencies in the well established characterization result of Kemeny (1959).

1 Introduction

Dissimilarity (and similarity) between two orders, rankings, queues, lists, etc. is typically measured by distance functions. A natural and frequently used way expressing dissimilarities is by counting the minimal number of swaps of consecutively ordered alternatives necessary to convert one order into the other. Its intuitive appeal may stem from the fact that these swaps cover the symmetric difference between the two orders. The reader may be familiar with the following (most probably non-exhaustive) list of names exhibiting this concept: the Kemeny (1959) distance, the Kendall-tau (1938) distance, swap distance, bubble-sort distance, inversion metric, word metric, permutation swap, the Hamming (1950) distance and so forth. According to Knuth (1998) prior to Kemeny (1959) and Kendall (1938), this simple idea can even be traced back to at least Cramer (1750). The axiomatic characterization of this distance function on weak orders can be found in Kemeny (1959) with proofs provided in Kemeny and Snell (1962). This paper analyzes measures based on this intuitive idea of swaps. We discuss two classes of such measures. In the first class, the swaps are weighted by their positions in the orders. In the second class, the swaps are weighted by the alternatives involved in these swaps.

We focus on an abstract setting of orders which spans all well-known “domains” of orders rather than addressing each individual domain separately. In this way, we bring about the essentials of the

*Department of Economics, School of Business and Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands; E-mail: b.can@maastrichtuniversity.nl, Phone number: +31 43 3883807. Authors are grateful to Bettina Klaus, Kristof Bosmans, and Uğur Özdemir for their comments on earlier drafts. The support of Netherlands Organisation for Scientific Research (NWO, Open Competitie) under the grant with project nr. 400-09-354, is gratefully acknowledged.

†*Corresponding author*: Department of Quantitative Economics, Maastricht University, P.O. Box 616, NL 6200 MD Maastricht, E-mail: t.storcken@ke.unimaas.nl, Phone number: +31 43 3883728, Fax number: +31 43 3884874

relations between the various conditions building our model. An order on a set of alternatives can be thought of as a ranking, list, tournament or a queue. Orders are commonly modeled as binary relations which somewhat compare alternatives in a context. Examples of such binary relations are, linear orders (also known as strict rankings), weak orders (rankings allowing ties), partial orders (incomplete rankings), tournaments and quasi orders¹. We distinguish between two types of domains in particular, “simple” and “aligned”. In a simple domain, an order may be converted into another by paths of swaps, each involving precisely two alternatives. For instance in the domain of linear orders, the swap between the rankings abc and bac involves only alternatives a and b , i.e., two consecutively ordered alternatives. In an aligned domain, however, such steps may involve more than two alternatives. For instance all three alternatives are involved in the conversion of the total indifference, (abc) , into the order, $(ab)c$, at which only the top two alternatives are indifferently ordered. Of course these conversions, such as from the former to the latter, are taken such that there are no other intermediate orders between them. Like in Storcken and de Swart (1992), these definitions on domains of orders capture all the well-known types including the aforementioned domains. We also use a broader definition of what a “measure” is, i.e., a function which assigns a non-negative real number to any two orders in a domain in a symmetric way. Like in Bogart (1973), we show that the crucial requirements of such measures are symmetry and non-negativity. This leads to a type of functions which are called “dissimilarity functions”. These so-called “dissimilarity functions” differ from typical metric (distance) functions by the two conditions, triangular inequality and identity of indiscernibles which are not imposed upon them.

In the original Kemeny (1959) characterization of the swap distance (from here onwards, we shall refer to it as the Kemeny distance), the *betweenness* axiom plays an important role. Betweenness requires that any intermediate order in the swap process of converting one order into another is on the “line segment” defined by the metric on these two orders. This implies that any path of elementary changes between two orders should additively sum up to the same dissimilarity. A weakening of this condition, such as *decomposability* introduced in Can (2014), requires only the existence of at least one such additive path. Together with other well-known axioms, these two leading axioms characterize two distinct classes of dissimilarity measures. Decomposability leads to a class of measures defined in terms of “positions” where orders differ, whereas betweenness leads to a class of measures defined in terms of the “pairs” on which orders differ. Via both approaches, we find out multiple characterizations of various measures.

Our first proposal in comparing orders is based on extending the idea of swapping according to the “positions” involved in these swaps. Consider two orders and assume the only swaps needed is at the top, this might cause more dissimilarity than at the bottom. Or conversely, swapping books at the top of a pile (a queue) may involve less effort than swapping two books at the bottom. Therefore we may want to weight swaps by the position of the alternatives taking part in them². We characterize the class of dissimilarity functions which we call “ τ -positional Kendall functions” by decomposability, neutrality and triangular inequality on all the aforementioned domains, simple and aligned, hence including the weak orders. This class allows for treatments that are sensitive to the positions of alternatives involved in the swaps.

¹A well-known example of quasi orders is an individual ordering some shades of grey, (not necessarily 50) and being indifferent in adjacent scales while ordering black over white strictly or vice versa. For further reading, see Luce (1956)

²The idea of weighting the swaps according to the positions can also be found in Can (2012, 2014) and Farnoud et al. (2012). Their results extend the Kemeny distance to a class of weighted distances on strict rankings only.

Our second proposal in comparing orders is based on extending the idea of swapping according to the “pairs” involved in these swaps. Consider two orders, such as two queues. If the only swap needed is between two grumpy old men, this might cause more dissimilarity than between two less strict people. Therefore we may want to weight swaps by the pair of alternatives taking part in them. This class of measures, i.e., “ ω -weighted Kemeny functions”, is characterized by *betweenness* alone. This class allows for treatments that are sensitive to the pairs of alternatives involved in the swaps. Additionally imposing *neutrality* yields the subclass “ α -scaled Kemeny functions” where these weights are all equal to a constant α . Further adding *normalization* restricts the class to the original Kemeny distance. Therefore a by-product of these results is the logical redundancy of the *reducibility* axiom introduced in the original characterization in Kemeny (1959), on all domains (simple and aligned) including weak orders. Aligned domains which are not simple, however, require special attention and a different proof method. A separate subsection is devoted to deal with these domains.

In each of the aforementioned classes, whether the weights are on positions or on pairs of alternatives, the leading axiom is either betweenness or a modification of it. To give an idea of the strength of this condition, consider any four orders R_{ab}, R'_{ab}, R_{ba} and R'_{ba} . In these four orders, alternatives a and b are ordered adjacently. Furthermore, R_{ba} is obtained from R_{ab} by reversing the strict order a above b to b over a at R_{ba} . R'_{ba} is obtained from R'_{ab} likewise. Betweenness imposes a circular distance structure on these four orders such that R_{ab} and R'_{ba} becomes antipodal (R'_{ab} and R_{ba} too). Therefore the arc (distance) between R_{ab} and R_{ba} equals to that of R'_{ab} and R'_{ba} . This alone yields a very vivid structure on the dissimilarity functions, i.e., swaps of identical alternatives have to be given the same weight (See Lemma 3). Consequently it shapes the characterization of the ω -weighted Kemeny functions whereas a modification of it, i.e. decomposability, shapes the characterization of the τ -weighted Kendall functions. We also discuss other weakenings of betweenness such as independence, and independence of pair conditions.

The paper is organized as follows. In Section 2 we introduce the abstract notion of orders and discuss some basic findings of it. In Section 3, we introduce dissimilarity functions together with the conditions mentioned above and show some logical inferences between some of these conditions. In Section 4 we characterize positional weighted dissimilarity functions and in Section 5 pairwise weighted dissimilarity functions. The latter section is split up in two subsections. The first subsection deals with the simple domains, whereas the second deals with the more general aligned domains which, among others, includes the set of weak orders. Section 6 elaborates on the logical independence of the various characterizing conditions, recaps the results and points out some possible directions for future investigation.

2 Basic concepts on orders

Let A denote a finite set of alternatives. Consider a binary relation R on A . As usual R is called *reflexive* if (x, x) is in R for all x in A . It is (*strongly*) *complete* if (x, y) is in R or (y, x) is in R for all alternatives x and y and it is *anti-symmetric* if (x, y) is not in R or (y, x) is not in R for all different alternatives x and y . For alternatives x and y and a relation R , we interpret as usual $(x, y) \in R$ as x is *ordered weakly above* y at R . In case (y, x) is also in R , we say that R *orders* x and y *indifferently*. In case (y, x) is not in R but (x, y) is in R , we say that x is *ordered strictly above* y . Given a relation R , we denote by $-R = \{(y, x) : (x, y) \in R\}$, the *reverse* of R .

For non-empty subsets B of A the *restriction* of R to B is denoted by $R_B = \{(x, y) \in B \times B : (x, y) \in R\}$. For a permutation π of A , the permuted relation R is defined by $\pi R = \{(\pi(x), \pi(y)) : (x, y) \in R\}$. Given two disjoint subsets X and Y of A , and relations R and R' on A the *concatenation* of the relations R_X with R'_Y is defined by $R_X \gg R'_Y = R_X \cup (X \times Y) \cup R'_Y$. When X or Y is a singleton set, e.g., $Y = \{y\}$, we abuse notation and write $R_X \gg y$. For relations R and R' on A the set symmetric difference operator \triangle is defined as usual: $R \triangle R' = (R \setminus R') \cup (R' \setminus R)$.

The characterizing results on dissimilarity functions hold on several domains of orders by identical proofs. Therefore we abstract from specific types of orders and introduce a general notion suiting our purpose. Like in Storcken and de Swart (1992), we define a domain of orders by some set of relations being closed under some intuitive appealing operations. Here, the operations are taken only as far as needed and adopted to the situation of a fixed set of alternatives.

Definition 1 A nonempty set of binary relations \mathcal{X} on A is called a domain (of orders) if \mathcal{X}

1. consists of reflexive relations only,
2. is closed under permutation; i.e. $\pi R \in \mathcal{X}$ for all $R \in \mathcal{X}$ and for all permutations π on A ,
3. is closed under linearization; i.e. $R_{A \setminus \{a\}} \gg a$ and $a \gg R_{A \setminus \{a\}}$ are in \mathcal{X} for all R in \mathcal{X} and for all a in A .

Given domain \mathcal{X} a *subdomain* \mathcal{Y} (of \mathcal{X}) is a subset of \mathcal{X} which is also a domain of orders.

Example 1 Note that by linearization we have that the set of all linear orders (complete, anti-symmetric, and transitive relations on A), say \mathcal{L} , is a subdomain of \mathcal{X} . It is straight forward to check that the set of weak orders (complete, transitive³ relations on A), the set of quasi-orders (complete, quasi-transitive⁴ relations on A), the set of partial orders (reflexive, anti-symmetric and transitive relations on A), and the set of tournaments (complete, anti-symmetric relations on A) are “domains” of orders. Moreover, if a domain \mathcal{X} contains a relation which is not anti-symmetric, say R , then there are different alternatives x and y such that both (x, y) and (y, x) are in R . Applying closedness under linearization and permutation repeatedly, we have that the following set:

$$\mathcal{L}^\cup = \{L^1 \cup L^2 : L^1, L^2 \in \mathcal{L} \text{ and } \exists x, y \in A, L^1 \setminus \{(x, y), (y, x)\} = L^2 \setminus \{(x, y), (y, x)\}\} \quad (1)$$

is a subset of \mathcal{X} . It is further straightforward to show that \mathcal{L}^\cup is a subdomain. In case a domain \mathcal{X} contains a relation which is not complete, say R , then there are different alternatives x and y such that both (x, y) and (y, x) are not in R . Then the following set of orders:

$$\mathcal{L}^\cap = \{L^1 \cap L^2 : L^1, L^2 \in \mathcal{L} \text{ and } \exists x, y \in A, L^1 \setminus \{(x, y), (y, x)\} = L^2 \setminus \{(x, y), (y, x)\}\} \quad (2)$$

is a subdomain of \mathcal{X} .

Consider orders R^1, R^2 and R^3 in domain \mathcal{X} . Order R^3 is said to be *between* R^1 and R^2 if $R^1 \cap R^2 \subseteq R^3 \subseteq R^1 \cup R^2$. Two different orders R^1, R^2 in \mathcal{X} form an *elementary change* (in \mathcal{X}) if there is no other order in \mathcal{X} , different from these two which is between them. In that case $R^1 \triangle R^2$ denotes the *type* of the elementary change $\{R^1, R^2\}$. Let \mathcal{E} denote the set of all (non ordered) tuples forming

³Transitivity means if (a, b) and (b, c) in R then (a, c) is in R .

⁴Quasi-transitivity means if (a, b) and (b, c) are in R and (b, a) and (c, b) are not, then (a, c) is in R and (c, a) is not.

an elementary change in \mathcal{X} , that is $\mathcal{E} = \{\{R^1, R^2\} : R^1, R^2 \in \mathcal{X} \text{ such that } R^1 \text{ and } R^2 \text{ form an elementary change in } \mathcal{X}\}$. Let the set of all possible types of elementary changes⁵ be denoted by $\mathcal{T} = \{R^1 \triangle R^2 : \{R^1, R^2\} \in \mathcal{E}\}$. Set \mathcal{E} and to some extent \mathcal{T} expresses the possible different changes in orders that may occur when altering one order into an other. In the sequel we only need the expressions in \mathcal{T} . We call domain \mathcal{X} *simple* if all elementary changes have similar types, that is if for any two elementary changes $\{R^1, R^2\}$ and $\{R^3, R^4\}$ there is a permutation π of A such that $\pi(R^1 \triangle R^2) = R^3 \triangle R^4$. Note that all domains in Example 1 except the domain of weak orders discussed are simple.

Two elementary changes $\{R^1, R^2\}$ and $\{R^3, R^4\}$ in \mathcal{E} *have the same position* if there exists a permutation π on A such that $\{R^3, R^4\} = \{\pi R^1, \pi R^2\}$. In that case we denote $\{R^1, R^2\} \sim \{R^3, R^4\}$. For explanation let $\pi R^1 = R^3$ and $\pi R^2 = R^4$ and further let R^2 be obtained from R^1 by reversing the preference between alternatives a and b only. Clearly the number of alternatives ordered above/below a at $R^1(R^2)$ is equal to that which is ordered above/below $\pi(a)$ at $R^3(R^4)$ and a same result holds for b . So, the “positions” of a at $R^1(R^2)$ and $\pi(a)$ at $R^3(R^4)$ are the same. This may motivate the definition on positions of elementary changes⁶. Note that having the same position is an equivalence relation on \mathcal{E} . Let $\tilde{\mathcal{E}}$ denote the set of equivalence classes in \mathcal{E} with respect to \sim . Further, for a tuple $\{R^1, R^2\}$ in \mathcal{E} , let $E[\{R^1, R^2\}]$ denote its equivalence class.

Example 2 Figure 2 depicts the elementary changes on the set of strict rankings (linear orders) where $A = \{a, b, c\}$. Here $abc = R^1$ means that at linear order R^1 , alternative a is ordered best b second best and c worst. Similar notations have likewise interpretations.

1. Based on their “positions”, elementary changes can be split in the following two groups. Group 1 consists of elementary changes: $\{R^1, R^2\}$, $\{R^3, R^6\}$, and $\{R^5, R^4\}$ with position “bottom”. Group 2 consists of elementary changes: $\{R^4, R^1\}$, $\{R^2, R^3\}$, and $\{R^6, R^5\}$ with position “top”.
2. Based on their “types”, elementary changes can be split in the following three groups. Group 1 consists of elementary changes $\{R^1, R^4\}$ and $\{R^3, R^6\}$ with type $\{(a, b), (b, a)\}$. Group 2 consists of elementary changes $\{R^2, R^3\}$ and $\{R^5, R^4\}$ with type $\{(a, c), (c, a)\}$. Group 3 consists of elementary changes $\{R^1, R^2\}$ and $\{R^5, R^6\}$ with type $\{(b, c), (c, b)\}$.

Let R^1, R^2, \dots, R^k be different orders in \mathcal{X} . The sequence $[R^1, R^2, R^3, \dots, R^k]$ is said to be a *path* from R^1 to R^k (in \mathcal{X}) if R^j is between R^i and R^l for all $1 \leq i < j < l \leq k$. As $R^i \cap R^l \subseteq R^j \subseteq R^i \cup R^l$ it follows that $(R^i \setminus R^j) \cup (R^j \setminus R^l) = R^i \setminus R^l$. Obviously $(R^i \setminus R^j)$ and $(R^j \setminus R^k)$ are disjoint. So, $R^1 \setminus R^2, R^2 \setminus R^3, R^3 \setminus R^4, \dots, R^{k-1} \setminus R^k$ partition $R^1 \setminus R^k$ with possibly empty parts. Similarly, $R^k \setminus R^{k-1}, R^{k-1} \setminus R^{k-2}, R^{k-2} \setminus R^{k-3}, \dots, R^2 \setminus R^1$ partition $R^k \setminus R^1$ with possibly empty parts. Noting that $R^i \neq R^j$ for all these possible $1 \leq i < j \leq k$ it follows $R^1 \triangle R^2, R^2 \triangle R^3, R^3 \triangle R^4, \dots, R^{k-1} \triangle R^k$ partition $R^1 \triangle R^k$.

Further, the path $[R^1, R^2, R^3, \dots, R^k]$ is called an *elementary path* if R^i and R^{i+1} form an elementary change for all $1 \leq i < k$. In that case $T[R^1, R^2, R^3, \dots, R^k]$ denotes the partition of $R^1 \triangle R^k$ in the types $R^1 \triangle R^2, R^2 \triangle R^3, R^3 \triangle R^4, \dots, R^{k-1} \triangle R^k$ of the elementary changes of that path from R^1 to R^k (with no empty parts). See Figure 2 for some examples of paths of elementary changes.

⁵See Example 2.2.

⁶See Example 2.1.

Remark 1 Consider two elementary paths $[R^1, R^2, R^3, \dots, R^k]$ and $[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$ both from orders R to R' in \mathcal{X} . We have that $T[R^1, R^2, R^3, \dots, R^k]$ and $T[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$ both partition $R \triangle R'$, i.e., the union of all types in each elementary path gives the symmetric set difference between R and R' . In general these partitions need not be equal. We say that \mathcal{X} is path independent if for all R and R' in \mathcal{X} and all elementary paths $[R^1, R^2, R^3, \dots, R^k]$ and $[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$ both from order R to R' in \mathcal{X} , the sets $T[R^1, R^2, R^3, \dots, R^k]$ and $T[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$ are equal. In that case we also write $T[R, \dots, R']$. The following Lemma shows that a simple domain is always path independent.

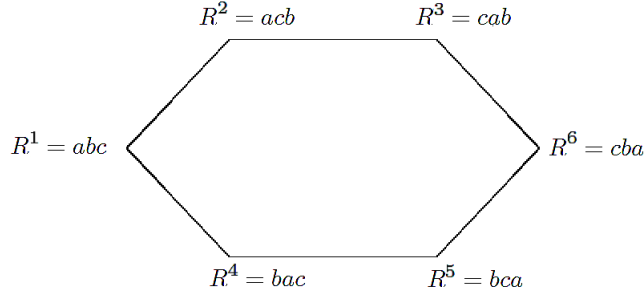


Figure 1: A graph for linear orders when $m = 3$. The two paths between R^1 and R^6 are $[R^1, R^2, R^3, R^6]$ and $[R^1, R^4, R^5, R^6]$.

Lemma 1 Let \mathcal{X} be a simple domain then \mathcal{X} is path independent. Further,

1. If all orders in \mathcal{X} are complete and anti-symmetric then for all types \mathbf{t} in \mathcal{T} there are different alternatives a and b such that $\mathbf{t} = \{(a, b), (b, a)\}$ and consequently $T[R, \dots, R'] = \{\{(a, b), (b, a)\} : (a, b) \in R \triangle R'\}$ for all different R and R' in \mathcal{X} ;
2. If some orders in \mathcal{X} are not complete or not anti-symmetric or both, then for all types \mathbf{t} in \mathcal{T} there are different alternatives a and b such that $\mathbf{t} = \{(a, b)\}$ and consequently $T[R, \dots, R'] = \{\{(a, b)\} : (a, b) \in R \triangle R'\}$ for all different R and R' in \mathcal{X} .

Proof. Let \mathcal{X} be a simple domain. Note that L^1 and L^2 in subdomain \mathcal{L} form an elementary change in \mathcal{L} whenever these differ on precisely one pair of alternatives, say a and b , being consecutively ordered in both these two linear orders. The type of elementary change $\{L^1, L^2\}$ is equal to $\{(a, b), (b, a)\}$. Further, let order R in \mathcal{X} be between L^1 and L^2 . As $L^1 \cap L^2 \subseteq R \subseteq L^1 \cup L^2$, $L^1 \cap L^2 = L^1 \setminus \{(a, b), (b, a)\}$ and $L^1 \cup L^2 = L^1 \cup \{(a, b), (b, a)\}$ it follows that $R \in \{L^1, L^2, L^1 \cup L^2, L^1 \cap L^2\}$. We may distinguish the following three cases.

Case 1 \mathcal{X} consists of complete and antisymmetric relations only. Consider an elementary change $\{L^1, L^2\}$ in \mathcal{L} of type $\{(a, b), (b, a)\}$. Let R in \mathcal{X} be between L^1 and L^2 . As $R \in \{L^1, L^2, L^1 \cup L^2, L^1 \cap L^2\}$ and $L^1 \cup L^2$ is not antisymmetric and $L^1 \cap L^2$ is not complete, R equals L^1 or L^2 . So, $\{L^1, L^2\}$ is an elementary change in \mathcal{X} . As \mathcal{X} is simple for all elementary changes there are different alternatives x and y such that the type of the elementary change equals $\{(x, y), (y, x)\}$. Take arbitrary orders

R and R' in \mathcal{X} and two elementary paths $[R^1, R^2, R^3, \dots, R^k]$ and $[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$ both from orders R to R' . It is sufficient to prove that $T[R^1, R^2, R^3, \dots, R^k] = T[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$. By the format of types of elementary changes we have $x^1, x^2, x^3, \dots, x^{k-1}, y^1, y^2, y^3, \dots, y^{k-1}$ such that elementary change $\{R^i, R^{i+1}\}$ has type $\{(x^i, y^i), (y^i, x^i)\}$ for $i \in \{1, 2, \dots, k-1\}$. So, $T[R^1, R^2, R^3, \dots, R^k] = \cup\{(x^i, y^i), (y^i, x^i) : i \in \{1, 2, \dots, k-1\}\}$. This partitions $R \triangle R'$. As R and R' are complete and anti-symmetric orders we have that $R \triangle R' = (R \setminus R') \cup (R' \setminus R)$ where $R' \setminus R = \{(x, y) : (y, x) \in R \setminus R'\}$. So, we may have chosen the x^i and y^i such that $\cup\{(x^i, y^i) : i \in \{1, 2, \dots, k-1\}\}$ partitions $(R \setminus R')$ and $\cup\{(y^i, x^i) : i \in \{1, 2, \dots, k-1\}\}$ partitions $(R' \setminus R)$. Similarly we have $\hat{x}^1, \hat{x}^2, \hat{x}^3, \dots, \hat{x}^{l-1}, \hat{y}^1, \hat{y}^2, \hat{y}^3, \dots, \hat{y}^{l-1}$ such that elementary change $\{\hat{R}^i, \hat{R}^{i+1}\}$ has type $\{(\hat{x}^i, \hat{y}^i), (\hat{y}^i, \hat{x}^i)\}$ for $i \in \{1, 2, \dots, l-1\}$ and $\cup\{(\hat{x}^i, \hat{y}^i) : i \in \{1, 2, \dots, l-1\}\}$ partitions $(R \setminus R')$ and $\cup\{(\hat{y}^i, \hat{x}^i) : i \in \{1, 2, \dots, l-1\}\}$ partitions $(R' \setminus R)$. Now it follows that $k = l$ and $T[R^1, R^2, R^3, \dots, R^k] = T[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$.

Case 2 Not all orders in \mathcal{X} are antisymmetric. As argued in Example 1 \mathcal{L}^\cup is a subdomain of \mathcal{X} . It follows that elementary changes in \mathcal{L}^\cup consist of a linear order say $L = L^1$ and a weak order say $W = L^1 \cup L^2$ where L^1 and L^2 form an elementary change in \mathcal{L} with type $\{(a, b), (b, a)\}$. In that case elementary changes in \mathcal{L}^\cup have a type $L \triangle W$ of the form $\{(x, y)\}$ for some different alternatives x and y . Because $L \triangle W$ is a singleton, an elementary change in \mathcal{L}^\cup is also an elementary change in \mathcal{X} . Because \mathcal{X} is simple all elementary changes in \mathcal{X} have a type which consists of precisely one ordered pair of different alternatives. By this it follows now straight forwardly that \mathcal{X} is path independent.

Case 3 Not all orders in \mathcal{X} are complete. Now \mathcal{L}^\cap is a subdomain of \mathcal{X} and elementary changes have types consisting of precisely one ordered pair of different alternatives. Hence, like before \mathcal{X} is path independent. ■

Let \mathcal{X} be a domain. Subdomain \mathcal{Y} is called *gapless* if every elementary change in \mathcal{Y} is an elementary change in \mathcal{X} . The following Lemma shows that any domain \mathcal{X} contains a gapless simple subdomain.

Lemma 2 *Let \mathcal{X} be a domain. Then there is a gapless simple subdomain \mathcal{S} .*

Proof. If all orders in \mathcal{X} are complete and antisymmetric, then we may take $\mathcal{S} = \mathcal{L}$. If all orders in \mathcal{X} are complete but not all are antisymmetric, then we may take $\mathcal{S} = \mathcal{L}^\cup$. If all orders in \mathcal{X} are antisymmetric but not all are complete, then we may take $\mathcal{S} = \mathcal{L}^\cap$. In case not all orders in \mathcal{X} are complete and not all are antisymmetric, then we may take $\mathcal{S} = \mathcal{L}^\cup \cup \mathcal{L}^\cap$. These choices (in the spirit of the proof of Lemma 1, and Example 1) satisfy the conditions on domains as well as that every elementary change in \mathcal{S} is also an elementary change in \mathcal{X} . ■

3 Dissimilarity functions

Let \mathcal{X} be a set of orders on A . Differences between orders in \mathcal{X} are indicated by a so called *dissimilarity* function, say δ . Although the main purpose of this function is to express dissimilarities between orders on A , in comparing our results to those of Kemeny (1959); Kemeny and Snell (1962) and in defining reducibility we consider it also to be a function on orders restricted to the subsets of A . It assigns to every pair of orders restricted to the same non empty subset of alternatives B , say R_B and R'_B , a *non-negative* real number indicated by $\delta(R_B, R'_B)$. Further, we assume that δ is *symmetric*, meaning $\delta(R_B, R'_B) = \delta(R'_B, R_B)$ for all orders R_B and R'_B . We like to point out

clearly here that extending δ to orders restricted to subsets of A is only needed to ease the above mentioned comparison. So, for the results discussed hereafter this extension is immaterial.

Next in relation with dissimilarity functions we define some well-known conditions for distance functions on orders. The first condition is a subcondition of any metric (distance) function.

Triangle Inequality: For all orders R, R' and R'' in \mathcal{X} ,

$$\delta(R, R'') \leq \delta(R, R') + \delta(R', R'').$$

The following four conditions are introduced by Kemeny (1959) and well-known. The condition of betweenness imposes that if an order is between two other orders, then triangle inequality holds with equality.

Betweenness⁷: For all orders R, R' and R'' in \mathcal{X} , such that R' is between R and R'' in \mathcal{X} ,

$$\delta(R, R'') = \delta(R, R') + \delta(R', R'').$$

Remark 2 *As also hinted out in Kemeny and Snell (1962), for any path of orders in \mathcal{X} , say $[R^1, \dots, R^l]$, betweenness implies $\delta(R^1, R^l) = \delta(R^1, R^2) + \delta(R^2, R^3) + \dots + \delta(R^{l-2}, R^{l-1}) + \delta(R^{l-1}, R^l)$. Using elementary paths this implies that under betweenness dissimilarities are completely determined by the dissimilarities between orders forming an elementary change.*

Neutrality defines that dissimilarities are not affected by to the actual naming of the alternatives, i.e., a relabeling of these alternatives should not change dissimilarity.

Neutrality: For all orders R and R' in \mathcal{X} and for any relabeling of alternatives $\pi : A \rightarrow A$,

$$\delta(R, R') = \delta(\pi R, \pi R').$$

We shall refer to the following condition as normalization. This is a condition which simply restricts the minimal non-zero dissimilarity to number one. Kemeny (1959) suggests this condition as a choice of “a unit of measurement”. Formally:

Normalization: The smallest non-zero dissimilarity between orders in \mathcal{X} equals 1.

At reducibility restrictions to subsets play an important role. It states that dissimilarity is not affected by deleting alternatives that are identically aligned at the top or bottom of the two orders. Formally:

Reducibility: For all orders R and R' in \mathcal{X} and any subset B of alternatives such that $R \triangle R' \subseteq B \times B$,

$$\delta(R, R') = \delta(R_B, R'_B).$$

The characterizations of classes of dissimilarity functions discussed hereafter are modeled on a fixed set of alternatives A . That is, subsets of A are all disregarded. To make comparison with Kemeny’s characterization more transparent, we introduce a condition which bears this independence idea of reducibility but does not consider dissimilarities on orders restricted to proper nonempty subsets⁸.

⁷Kemeny (1959) also requires the reverse, i.e. if equality holds then R' is between R and R'' . As this is not needed here we do not impose it.

⁸That is to say, the reducibility by construction has to be defined on a varying set of alternatives whereas independence does not require such a scheme. Therefore under the same setting independence is weaker than reducibility.

Independence: For pairs of orders R, R' and \hat{R}, \hat{R}' in \mathcal{X} and any nonempty subsets B of A , such that $R_B = \hat{R}_B$, $R'_B = \hat{R}'_B$ and $R \triangle R' \subseteq B \times B$ and $\hat{R} \triangle \hat{R}' \subseteq B \times B$

$$\delta(R, R') = \delta(\hat{R}, \hat{R}').$$

Example 3 Consider $R = \text{abcxyzdef}$, $R' = \text{abczyxdef}$ and $\hat{R} = \text{cbxyzfaed}$, $\hat{R}' = \text{cbzyxfaed}$. Then $R_B = \hat{R}_B$, $R'_B = \hat{R}'_B$ and $R \triangle R' \subseteq B \times B$ and $\hat{R} \triangle \hat{R}' \subseteq B \times B$ where $B = \{x, y, z\}$. Independence requires that $\delta(R, R') = \delta(\hat{R}, \hat{R}')$.

As reducibility implies that $\delta(R, R') = \delta(R_B, R'_B)$ and $\delta(\hat{R}, \hat{R}') = \delta(\hat{R}_B, \hat{R}'_B)$, and because of $R_B = \hat{R}_B$, $R'_B = \hat{R}'_B$, it follows immediately that reducibility implies independence. In fact, in Farnoud et al. (2012), Kemeny's reducibility is explained in terms of what here is called independence.

The following property originates from Bogart⁹ (1973). It is implied by betweenness.

Independence of pair addition: For pairs of orders R, R' and \hat{R}, \hat{R}' in \mathcal{X} and alternatives a and b such that $R_B \setminus R'_B = \hat{R}_B \setminus \hat{R}'_B = \{(a, b)\}$ and $R'_B \setminus R_B = \hat{R}'_B \setminus \hat{R}_B = \emptyset$

$$\delta(R, R') = \delta(\hat{R}, \hat{R}').$$

The following lemma is used to show that betweenness implies independence and independence of pair addition. It is best illustrated by four points, say a, \hat{a}, b and \hat{b} on a circuit where a is antipodal with \hat{a} and b with \hat{b} . Considering only paths along this circuit it follows that both a and \hat{a} are between b and \hat{b} and vice versa, albeit via different paths. In such a situation it follows that the distance between a and b is equal to that between \hat{a} and \hat{b} . The Lemma now shows that this holds for any such four points satisfying this mutual betweenness. We thereafter provide a figure to illustrate the lemma.

Lemma 3 Let R, R', \hat{R} and \hat{R}' be orders in domain \mathcal{X} such that i) both R and \hat{R}' are between R' and \hat{R} and ii) both R' and \hat{R} are between R and \hat{R}' . Let δ be a dissimilarity function which satisfies betweenness. Then

$$\delta(R, R') = \delta(\hat{R}, \hat{R}')$$

Proof. As δ satisfies betweenness we have

$$\begin{aligned} \delta(R', R) + \delta(R, \hat{R}) &= \delta(R', \hat{R}') + \delta(\hat{R}', \hat{R}) \text{ and} \\ \delta(R, R') + \delta(R', \hat{R}') &= \delta(R, \hat{R}) + \delta(\hat{R}, \hat{R}'). \end{aligned}$$

Adding these two equations applying symmetry and further simplification yields the desired equality. ■ Using Lemma 3, we can now show that betweenness implies independence.

Lemma 4 Let δ be a dissimilarity function on domain \mathcal{X} . Let δ satisfy betweenness. Let R, R' and \hat{R}, \hat{R}' be orders in \mathcal{X} such that $R \setminus R' = \hat{R} \setminus \hat{R}'$ and $\hat{R}' \setminus \hat{R} = R' \setminus R$. Then $\delta(R, R') = \delta(\hat{R}, \hat{R}')$. Moreover, δ satisfies independence and independence of pair addition.

⁹Bogart calls this condition D5.

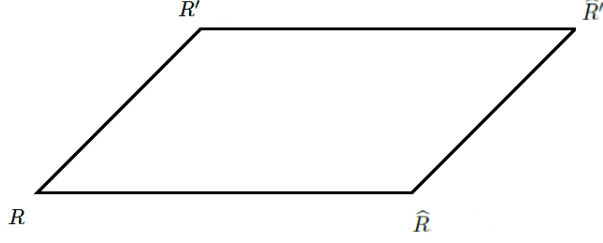


Figure 2: R', R, \widehat{R} , and \widehat{R}' are as defined in Lemma 3. Betweenness hence implies that $\delta(R, R') = \delta(\widehat{R}, \widehat{R}')$.

Proof. In view of Lemma 3 it is sufficient to prove that R and \widehat{R}' are both between R' and \widehat{R} , and that R' and \widehat{R} are both between R and \widehat{R}' . This boils down to prove each of the following four statements.

$$\begin{aligned} R' \cap \widehat{R} &\subseteq R \subseteq R' \cup \widehat{R}. \\ R' \cap \widehat{R} &\subseteq \widehat{R}' \subseteq R' \cup \widehat{R}. \\ R \cap \widehat{R}' &\subseteq R' \subseteq R \cup \widehat{R}'. \\ R \cap \widehat{R}' &\subseteq \widehat{R} \subseteq R \cup \widehat{R}'. \end{aligned}$$

We only prove the first one as the other three follow similarly. As $R' \setminus R \subseteq \widehat{R}' \setminus \widehat{R}$, we have $(R' \cap \widehat{R}) \subseteq R$. Because of $R \setminus R' \subseteq \widehat{R} \setminus \widehat{R}'$ it follows that $R \subseteq R' \cup \widehat{R}$.

(Moreover part) Consider orders R, R' and $\widehat{R}, \widehat{R}'$ in \mathcal{X} to which either independence or independence of pair addition applies. Now obviously $R \setminus R' = \widehat{R} \setminus \widehat{R}'$ and $\widehat{R}' \setminus \widehat{R} = R' \setminus R$ and we are done by the previous argument. ■

Remark 3 For dissimilarity functions δ on simple domains \mathcal{X} satisfying betweenness Lemma 4 implies that if R, R' and $\widehat{R}, \widehat{R}'$ are elementary changes of the same type, then $\delta(R, R') = \delta(\widehat{R}, \widehat{R}')$. Indeed after possible renaming the orders on such domains there are alternatives a and b such that $R \setminus R' = \widehat{R} \setminus \widehat{R}' = \{(a, b)\}$ and either $R' \setminus R = \widehat{R}' \setminus \widehat{R} = \{(b, a)\}$ or $R' \setminus R = \widehat{R}' \setminus \widehat{R} = \emptyset$. So, Lemma 4 yields $\delta(R, R') = \delta(\widehat{R}, \widehat{R}')$.

The following condition is introduced in Can (2012, 2014) as a relaxation of betweenness (also in Farnoud and Milenkovic (2012)). Instead of demanding that all paths between two orders have the same minimal length, it only requires that there is at least one such path. That is, if two orders do not form an elementary change then we can find a third which is on a shortest path between the former two. In that way, at least one shortest elementary path between these orders exists. Hence a dissimilarity is decomposable into a sequence of dissimilarities on elementary changes.

Decomposability: For all orders R, R' which do not form an elementary change in \mathcal{X} , there exist R'' such that

$$\delta(R, R') = \delta(R, R'') + \delta(R'', R').$$

The following Lemma shows that on some simple domains decomposability and independence together are equivalent to betweenness.

Lemma 5 *Let \mathcal{X} be a simple domain. Let δ be a dissimilarity function on \mathcal{X} .*

1. *If all orders in \mathcal{X} are complete and δ is decomposable and independent, then δ satisfies betweenness;*
2. *If all orders in \mathcal{X} are anti-symmetric and δ is decomposable and independent, then δ satisfies betweenness;*
3. *If some orders in \mathcal{X} are not complete and δ is decomposable and independent of pair addition, then δ satisfies betweenness;*
4. *If some orders in \mathcal{X} are not anti-symmetric and δ is decomposable and independent of pair addition, then δ satisfies betweenness.*

Proof. Let δ be independent and decomposable. Consider R and R' in \mathcal{X} . To prove betweenness it is sufficient to argue that all paths between R and R' have equal length $\delta(R, R')$. Because \mathcal{X} is a simple domain by Lemma 1 it is path independent. So, for any two paths from R to R' say $[R^1, R^2, R^3, \dots, R^k]$ and $[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$ we have $k = l$ and the type sets $T[R^1, R^2, R^3, \dots, R^k]$ are equal $T[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$.

(Proof of 1) As all orders \mathcal{X} are complete it follows by Lemma 1 that for two elementary changes say $\{R^1, R^2\}$ and $\{\hat{R}^i, \hat{R}^{i+1}\}$ with equal type there are alternatives a and b such that $R_{\{a,b\}}^1 = \hat{R}_{\{a,b\}}^i$, $R_{\{a,b\}}^2 = \hat{R}_{\{a,b\}}^{i+1}$, $R^1 \triangle R^2 \subseteq \{a, b\} \times \{a, b\}$ and $\hat{R}^i \triangle \hat{R}^{i+1} \subseteq \{a, b\} \times \{a, b\}$. So, independence implies that $\delta(R^1, R^2) = \delta(\hat{R}^i, \hat{R}^{i+1})$. Because $T[R^1, R^2, R^3, \dots, R^k]$ equals $T[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$,

this yields that $\sum_{i=1}^{k-1} \delta(R^i, R^{i+1}) = \sum_{i=1}^{k-1} \delta(\hat{R}^i, \hat{R}^{i+1})$. By decomposability we may assume without loss of generality that $\delta(R, R') = \sum_{i=1}^{k-1} \delta(R^i, R^{i+1})$. So, $\delta(R, R') = \sum_{i=1}^{k-1} \delta(\hat{R}^i, \hat{R}^{i+1})$ for any path

$[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^k]$ from R to R' .

(Proof of 2) Is similar to (1).

(Proof of 3) As some orders \mathcal{X} are not complete it follows by Lemma 1 that for two elementary changes say $\{R^1, R^2\}$ and $\{\hat{R}^i, \hat{R}^{i+1}\}$ with equal type there are alternatives a and b such that $R_{\{a,b\}}^s \setminus R_{\{a,b\}}^t = \hat{R}_{\{a,b\}}^{\hat{s}} \setminus \hat{R}_{\{a,b\}}^{\hat{t}} = \{(a, b)\}$, $R_{\{a,b\}}^t \setminus R_{\{a,b\}}^s = \hat{R}_{\{a,b\}}^{\hat{t}} \setminus \hat{R}_{\{a,b\}}^{\hat{s}} = \emptyset$ for some renumbering s and t of $\{1, 2\}$ and \hat{s} and \hat{t} of $\{i, i+1\}$. But then independence of pair addition implies that $\delta(R^1, R^2) = \delta(\hat{R}^i, \hat{R}^{i+1})$ and because $T[R^1, R^2, R^3, \dots, R^k]$ equals $T[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$, this yields

that $\sum_{i=1}^{k-1} \delta(R^i, R^{i+1}) = \sum_{i=1}^{k-1} \delta(\hat{R}^i, \hat{R}^{i+1})$. By decomposability we may assume without loss of general-

ity that $\delta(R, R') = \sum_{i=1}^{k-1} \delta(R^i, R^{i+1})$. So, $\delta(R, R') = \sum_{i=1}^{k-1} \delta(\hat{R}^i, \hat{R}^{i+1})$ for any path $[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^k]$

from R to R' .

(Proof of 4) Is similar to (3). ■

Considering Lemma 5 in relation with independence of pair addition we see that part 3 and 4 do not cover the intersection of parts 1 and 2. Indeed on simple domains which consist of complete and anti-symmetric orders independence of pair addition has no bite. Because on such domains elementary changes consist of the addition and removal of one pair. Considering the same Lemma in relation with independence we see that part 1 and 2 do not cover the intersection of part 3 and 4. On simple domains at which there exist incomplete orders as well as not anti-symmetric orders for all elementary changes there are alternatives a and b such that these have type $\{(a, b)\}$. But now there are elementary changes $\{R, R'\}$ and $\{\widehat{R}, \widehat{R}'\}$ both of type $\{(a, b)\}$, but $(a, b), (b, a) \in R$, $(b, a) \in R'$ and $(a, b) \notin R'$, $(a, b) \in \widehat{R}$ and $(b, a) \notin \widehat{R}$, and $(a, b) \notin \widehat{R}'$ and $(b, a) \notin \widehat{R}'$. So, at these changes independence has no bite.

4 Comparing orders by positional weighting

Here dissimilarities between orders are based on the positions at which swaps or mergers, to convert the one order into the other, take place. For instance swapping two wagons in the middle of a train may require more effort than a swap of first two (or last two) wagons.

It is not difficult to see that dissimilarity based on the positions of swaps may violate betweenness. Indeed consider the two paths in Figure 2. As mentioned in Remark 2, betweenness requires that they both lead to the same sum. However, the first path requires a consecutive swap pattern at bottom, top, bottom positions whereas the second path requires a pattern at top, bottom, top. Hence, betweenness is only satisfied if weights at top and bottom position equal.

Weighting the position at which swaps or mergers in elementary changes occur is now formalized by a *positional weight* function τ . Function τ assigns to every possible position in $\underline{\mathcal{E}}$ a non-negative number. Given an elementary change $\{R^1, R^2\}$ the number $\tau(E[\{R^1, R^2\}])$ assigned by τ to the equivalence class of $\{R^1, R^2\}$ can be interpreted as the relative effort to exercise an elementary change at position $E[\{R^1, R^2\}]$. The τ -positional Kendall distance function on \mathcal{X} is defined for all orders¹⁰ R and R' by

$$\kappa^\tau(R, R') = \min \left\{ \sum_{i=1}^l \tau(E[\{R^i, R^{i+1}\}]) : [R^1, \dots, R^{l+1}] \text{ is a path of elementary changes with } l \geq 1 \text{ from } R \text{ to } R' \right\} \quad (3)$$

So, τ -positional Kendall distance minimizes the sum of positional weights over all paths of elementary changes in \mathcal{X} between R and R' .

Next, we present the results on τ -positional Kendall distances. Theorem 1 below can also be found (though only on the domain of linear orders) in Farnoud and Milenkovic (2012), and in Can (2012, 2014). Hence, Theorem 1 can be seen as an extension of these to any domain of orders on A .

Theorem 1 *Let \mathcal{X} be a set of orders on A . Let δ be a dissimilarity function on \mathcal{X} . Then δ satisfies decomposability, neutrality and triangle inequality if and only if there exists a positional weight function τ such that $\delta = \kappa^\tau$.*

¹⁰Note that when $R = R'$, we abuse notation for simplicity and assume $\kappa^\tau(R, R') = 0$ instead of defining pairs of identical orders as a null elementary change.

Proof. The proof of the if-part is straight forward and left to the reader.

(only if-part) Let δ be decomposable and neutral. Neutrality implies that δ assigns equal values to any two elementary changes which are at the same position. Therefore we can define a function τ from $\frac{\mathcal{E}}{\sim}$ to the set of non-negative real numbers \mathbb{R}_+ by $\tau(E[\{R^1, R^2\}]) = \delta(R^1, R^2)$ where $E[\{R^1, R^2\}]$ is an arbitrary chosen equivalence class in $\frac{\mathcal{E}}{\sim}$ represented by elementary change $\{R^1, R^2\}$. Decomposability now implies that for any two orders R and R' there is a path of elementary changes, say $[R^1, \dots, R^k]$ from R to R' such that

$$\delta(R, R') = \sum_{i=1}^{k-1} \tau(E[\{R^i, R^{i+1}\}]).$$

Triangle inequality implies that $\delta(R, R')$ equals the minimum over all such paths of elementary changes from R to R' . Hence, $\delta(R, R')$ equals the summation at Equation 3 and therewith $\delta(R, R') = \kappa^\tau(R, R')$. ■

Below we would like to demonstrate a τ -positional Kendall distance on the domain of weak orders.

Example 4 Let $A = \{a, b, c\}$. Consider the domain of weak orders in Figure 4 and let us categorize them as i) strict orders, S , $(R^1, R^2, R^3, R^4, R^5, R^6)$, ii) weak orders with indifference at top, IT , (R^{12}, R^{56}, R^{23}) , iii) weak orders with indifference at bottom, IB , (R^{12}, R^{36}, R^{45}) , and iv) full indifference, FI , (R^0) . There are 4 different equivalence classes regarding the “positions” of elementary changes and $\frac{\mathcal{E}}{\sim} = \{E(S, IT), E(S, IB), E(FI, IT), E(FI, IB)\}$. Let τ assign the following weights to these positions; $\tau(S, IT) = 3$, $\tau(S, IB) = 2$, $\tau(FI, IT) = 4$, $\tau(FI, IB) = 4.5$. Then the distance κ_τ between each elementary change is depicted below.

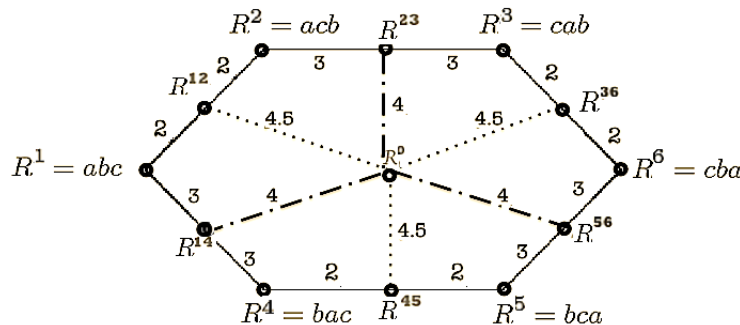


Figure 3: A graph for weak orders when $m = 3$. Given the τ weights in Example 4, decomposability necessitates that $\kappa^\tau(R^1, R^0) = 2 + 4.5 = 6.5$, $\kappa^\tau(R^1, R^{14}) = 3$, $\kappa^\tau(R^1, R^{12}) = 2$, $\kappa^\tau(R^{14}, R^0) = 4$, $\kappa^\tau(R^{12}, R^0) = 4.5$, $\kappa^\tau(R^1, R^6) = 2 + 4.5 + 4.5 + 2 = 13$.

This weight distribution defines all the distances between any two weak orders due to decomposability. Note that the domain of weak orders is a superset of \mathcal{L}^\cup which is mentioned in Example 1 since $\mathcal{L}^\cup = S \cup IT \cup IB$. Hence the domain of weak orders $\mathcal{X} = \mathcal{L}^\cup \cup FI$.

5 Comparing orders by pairwise weighting

Dissimilarities between orders may be based on their symmetric difference as this expresses the pairs of alternatives which have to be swapped to change one order into the other. Swapping different pairs of alternatives may invoke different amounts of effort. Therefore we may want to weight swaps by the pair of alternatives taking part in it. For instance, swapping two consecutively ordered tourists in a service queue may require less tact and delicacy than swapping two natives in the same queue. Introducing weights on pairs of alternatives enables to treat each pair differently. First we study pairwise weighted dissimilarities on simple domains and in a second subsection we treat more general domains including the domain of weak orders.

5.1 Simple domains

Assume that \mathcal{X} is a simple domain. So, by Lemma 1 it is path independent. We introduce a *weight function* ω from the set of all types \mathcal{T} of elementary changes in \mathcal{X} to \mathbb{R}_+ . Based on such a weight function, for any two orders $R, \hat{R} \in \mathcal{X}$, we define ω -*weighted Kemeny distance* on \mathcal{X} by

$$\kappa_\omega(R, \hat{R}) = \sum_{t \in T[R, \dots, \hat{R}]} \omega(t). \quad (4)$$

A ω -weighted Kemeny distance sums the weights of the pairs to be swapped or merged to obtain one order from another. Interpreting these weights as the relative effort to execute such swaps or mergers, this distance then measures the total effort to build R from \hat{R} by swapping and merging pairs of alternatives.

Equalizing all weights to a scalar, say α , means that the distance is only based on the number of pairs that have to be swapped or merged. This defines an α -*scaled Kemeny distance* on \mathcal{X} by

$$\kappa_\alpha(R, \hat{R}) = \alpha \cdot \#T[R, \dots, \hat{R}]. \quad (5)$$

Finally, the Kemeny distance is obtained by imposing the condition of normalization on this class of α -scaled Kemeny distances. Indeed let the smallest positive distance be equal to 1. For an α -scaled Kemeny distances this implies that $\alpha = 1$. For clear reference we define the *Kemeny distance* κ on \mathcal{X} by:

$$\kappa(R, \hat{R}) = \#T[R, \dots, \hat{R}]. \quad (6)$$

By varying ω and α we obtain a class of respectively *weighted Kemeny* and *scaled Kemeny* distances. Obviously, Kemeny distance is a scaled Kemeny distance and every scaled Kemeny distance is a weighted Kemeny distance. Clearly all weighted Kemeny distance functions satisfy betweenness.

Besides characterizing the class of weighted Kemeny distances on simple domains the following theorem shows that for these distance functions neutrality together with the condition of betweenness characterizes the class of scaled Kemeny distance functions. And neutrality together with the conditions of betweenness and normalization characterize the Kemeny distance function.

Theorem 2 *Let \mathcal{X} be a simple domain. Let δ be a dissimilarity function on \mathcal{X} . Then*

1. δ satisfies betweenness if and only if there exists a weight function ω such that $\delta = \kappa_\omega$,

2. δ satisfies betweenness and neutrality if and only if there exists a scalar $\alpha > 0$ such that $\delta = \kappa_\alpha$,

3. δ satisfies betweenness, neutrality, and normalization if and only if $\delta = \kappa$.

Proof. The proofs of each of the three *if-parts* are straight forward and left to the reader.

(*Only-if-part of 1*) Let δ satisfy the condition of betweenness. By Remark 3 for any two elementary changes R and R' on the one hand and \bar{R} and \bar{R}' on the other of the same type we have that $\delta(R, R') = \delta(\bar{R}, \bar{R}')$. Therefore we may define a weight function ω which assigns an arbitrary type \mathbf{t} in \mathcal{T} weight $\omega(\mathbf{t}) = \delta(R, R')$, where R and R' form an elementary change in type \mathbf{t} . For an arbitrary pair of relations say R^1 and R^2 in \mathcal{X} we now may conclude by Remark 2

$$\begin{aligned}\delta(R^1, R^2) &= \sum_{\mathbf{t} \in T[R^1, \dots, R^2]} \omega(\mathbf{t}) \\ &= \kappa_\omega(R^1, R^2).\end{aligned}$$

(*Only-if-part of 2*) By (1) we have that $\delta = \kappa_\omega$. Neutrality implies that $\omega(\mathbf{t}^1) = \omega(\mathbf{t}^2)$ for all types $\mathbf{t}^1, \mathbf{t}^2 \in \mathcal{T}$. Taking $\alpha = \omega(\mathbf{t})$ for some $\mathbf{t} \in \mathcal{T}$, it follows for an arbitrary pair of preferences say R^1 and R^2 in \mathcal{X} that

$$\begin{aligned}\delta(R^1, R^2) &= \sum_{\mathbf{t} \in T[R^1, \dots, R^2]} \omega(\mathbf{t}) \\ &= \alpha \cdot \#T[R^1, \dots, R^2] \\ &= \kappa_\alpha(R^1, R^2).\end{aligned}$$

So, $\delta = \kappa_\alpha$.

(*Only-if-part of 3*) By (2) we have $\delta = \kappa_\alpha$ and by normalization it follows that $\alpha = 1$. So, $\delta = \kappa$.

The theorem above holds for all domains of orders that are simple including linear orders, quasi orders, partial orders, tournaments and many more. Hence, the above theorem is closely related to Kemeny's characterization of the Kemeny distance. One obvious difference is that the set of alternatives, A , is fixed in the theorem above¹¹ whereas it is not in Kemeny (1959); Kemeny and Snell (1962). Another difference is that Kemeny's result considers the larger domain of weak orders. This domain also satisfies the conditions in Definition 1 but it is not simple. In the following subsection we discuss more general domains, including the set of weak orders, for which a similar result holds. By Theorem 2.1 other well-known results, such as Theorem 2 in Bogart (1973), can also be deduced¹².

Note that in the characterization presented above reducibility is not a characterizing condition whereas it is in Kemeny (1959) and in Kemeny and Snell (1962). Reducibility appears in the form of independence, as a consequence of betweenness, which enables to base the characterization of Kemeny distance on a fixed set without considering its subsets or supersets¹³.

¹¹Note that if the axioms are defined in the style of Kemeny (1959), i.e., defined also on restrictions of orders to each subset of alternatives, then the results still hold since our axioms are weaker in this sense.

¹²Considering dissimilarity functions δ which satisfy betweenness we have that $\delta = \kappa_\omega$. Adding condition D6 of Bogart (1973), we have that all weights of all elementary changes are equal. So, δ a scaled Kemeny distance function. Finally, condition D7 of Bogart (1973) is just the normalisation condition. So, adding this yields that $\delta = \kappa$.

¹³An independent re-characterization of the Kemeny distance on weak orders can be found in Can and Storcken (2013).

Corollary 1 paralleling Theorem 2 follows obviously from it and Lemma 5 and is therefore stated without proof.

Corollary 1 *Let \mathcal{X} be a simple domain. Let δ be a dissimilarity function on \mathcal{X} .*

A *Let all orders in \mathcal{X} be complete or all orders in \mathcal{X} be anti symmetric. Then*

1. *δ satisfies independence and decomposability if and only if, there is a weight function ω such that $\delta = \kappa_\omega$.*
2. *δ satisfies independence, decomposability and neutrality if and only if, there is a scalar α such that $\delta = \kappa_\alpha$.*
3. *δ satisfies independence, decomposability, neutrality and normalization if and only if, $\delta = \kappa$.*

B *Let some orders in \mathcal{X} be not complete or some orders in \mathcal{X} be not anti-symmetric. Then*

1. *δ satisfies independence of pair addition and decomposability if and only if, there is a weight function ω such that $\delta = \kappa_\omega$.*
2. *δ satisfies independence of pair addition, decomposability and neutrality if and only if, there is a scalar α such that $\delta = \kappa_\alpha$.*
3. *δ satisfies independence of pair addition, decomposability, neutrality and normalization if and only if, $\delta = \kappa$.*

Note that Corollary 1 essentially uses the simplicity of the domain. On such domains, Lemma 5 shows that decomposability together with independence (or independence of pair addition) implies betweenness. Therefore, Theorem 2 gives Corollary 1. The example below shows that neither Lemma 5 nor Corollary 1 can be extended to the domain of weak orders.

Example 5 *Consider the set of all weak orders on $A = \{a, b, c\}$. Let δ be equivalent to κ except for the cases where $\{R, R'\} = \{a(bc), (abc)\}$ or $\{R, R'\} = \{(bc)a, (abc)\}$. For those cases let $\delta(R, R') = 3$. Clearly, δ is a dissimilarity function which is not equal to κ . It obviously satisfies decomposability, normalization and triangle inequality. It evidently does not satisfy neutrality and betweenness. However, it satisfies “independence” and “independence of pair addition”. Note that on elementary changes having a single ordered pair as “type”, both independence conditions hold by definition. On elementary changes with types consisting of multiple ordered pairs, e.g. $R = a(bc)$ and $R' = (abc)$, they hold as well. In such cases, i.e., from R to R' , all three alternatives are involved in the change, and the swap requires two ordered pairs to be added, i.e., $\{(b, a), (c, a)\}$. Therefore both independence conditions lose their bite and the characterization of κ above does not work for this non-simple domain.*

5.2 Aligned domains

Theorem 2 shows that dissimilarity functions satisfy betweenness only these are weighted Kemeny distances. That is for simple domains. Here we extend this result to the so-called *aligned* domains. Among these is the domain of weak orders.

Definition 2 Consider a domain \mathcal{X} and a gapless simple subdomain \mathcal{S} of \mathcal{X} . \mathcal{X} is an aligned domain (towards \mathcal{S}) if for any R and R' in \mathcal{X} , there exists $S \in \mathcal{S}$ such that:

1. R is between R' and S or,
2. R' is between R and S .

If a domain \mathcal{X} is an aligned domain, we say that all orders in \mathcal{X} are *aligned towards* simple domain \mathcal{S} . Although a simple domain is aligned towards itself, a non-simple domain does not have to be aligned to any of its simple subdomains (see Example 6). Note that the domain of weak orders is an aligned domain (in particular towards the gapless simple subdomain \mathcal{L}^\cup).

Note that in a simple domain either i) all types are singleton, e.g., $\{(a, b)\}$, or ii) all types are doubleton (an ordered pair and its reverse), e.g., $\{(a, b), (b, a)\}$. In an aligned domain, however, the types may vary depending on the paths taken between two orders, say R and R' . Therefore we cannot use the Equations 4, 5, and 6 by assigning weights to types directly. However, as explained in Remark 1, the types in each path, e.g., $[R^1, R^2, R^3, \dots, R^k]$ and $[\hat{R}^1, \hat{R}^2, \hat{R}^3, \dots, \hat{R}^l]$ both from orders R to R' partition the symmetric set difference between the two orders, i.e., $R \triangle R'$. Therefore an equivalent definition of the aforementioned Equations can be provided by assigning weights to the ordered pairs in the symmetric set difference. Below we provide these straightforward extensions of κ_ω , κ_α , and κ (in Equations 4, 5, and 6) to aligned domains (in Equations 4*, 5*, and 6*).

We introduce a *weight function* v from the set of all ordered pairs with distinct alternatives to \mathbb{R}_+ , i.e., for any two different alternatives $a, b \in A$. Based on such a weight function, for any two orders $R, \hat{R} \in \mathcal{X}$, we define *v-weighted Kemeny distance* on aligned domain \mathcal{X} by:

$$\kappa_v^*(R, \hat{R}) = \sum_{(a,b) \in R \triangle \hat{R}} v(\{(a, b)\}) \quad (4^*)$$

Equalizing all weights to a scalar, say α , defines an α -scaled *Kemeny distance* on aligned domain \mathcal{X} :

$$\kappa_\alpha^*(R, \hat{R}) = \alpha \cdot \#(R \triangle \hat{R}) \quad (5^*)$$

Finally, the Kemeny distance is obtained by imposing the condition of normalization on this class of α -scaled Kemeny distances.

$$\kappa_\alpha^*(R, \hat{R}) = \#(R \triangle \hat{R}) \quad (6^*)$$

Next, we prove that Theorem 2 extends to all aligned domains \mathcal{X} . That is on such domains a dissimilarity function δ satisfies betweenness if and only if, $\delta(R, R') = \kappa_v^*$ for all orders R and R' . The following two Lemma's are needed to ease the deduction of the statement spelled out above.

Lemma 6 Let δ be a dissimilarity function on domain \mathcal{X} satisfying betweenness. Let R, R^1 and R^2 be three orders in \mathcal{X} . Let $\delta(R, R^1) = \kappa_v^*(R, R^1)$ and $\delta(R^1, R^2) = \kappa_v^*(R^1, R^2)$. Let R^2 be between R and R^1 or R^1 be between R and R^2 . Then $\delta(R, R^2) = \kappa_v^*(R, R^2)$.

Proof. In case R^1 be between R and R^2 the result follows by betweenness. Suppose R^2 be between R and R^1 . Then betweenness implies

$$\delta(R, R^1) = \delta(R, R^2) + \delta(R^2, R^1).$$

So,

$$\begin{aligned}\delta(R, R^2) &= \delta(R, R^1) - \delta(R^2, R^1) \\ &= \kappa_v^*(R, R^1) - \kappa_v^*(R^2, R^1).\end{aligned}$$

Using the definition of κ_v^* this leads to

$$\begin{aligned}\delta(R, R^2) &= \sum_{(a,b) \in R \triangle R^1} v(\{(a,b)\}) - \sum_{(a,b) \in R^1 \triangle R^2} v(\{(a,b)\}) \\ &= \sum_{(a,b) \in R \triangle R^2} v(\{(a,b)\}) \\ &= \kappa_v^*(R, R^2).\end{aligned}$$

■

Lemma 7 *Let \mathcal{X} be a domain with gapless simple subdomain \mathcal{S} . Let δ be a dissimilarity function on \mathcal{X} satisfying betweenness. Let $R \in \mathcal{X}$ and $S^1, S^2 \in \mathcal{S}$. If $\delta(R, S^1) = \kappa_v^*(R, S^1)$, then $\delta(R, S^2) = \kappa_v^*(R, S^2)$.*

Proof. In the view of Lemma 6, it is sufficient to prove that S^1 is between S^2 and R , or S^2 is between S^1 and R . Without loss of generality we may assume that S^1 and S^2 form an elementary change in \mathcal{X} . We distinguish the following two cases:

Case 1: \mathcal{X} consists of antisymmetric and complete relations only. Then the type of elementary changes in the gapless simple subdomain \mathcal{S} consists of two ordered pairs. Hence we may assume $S^1 = (S^2 \setminus \{(y, x)\}) \cup \{(x, y)\}$ for some $x, y \in A$. Then $S^1 \triangle S^2 = \{(x, y), (y, x)\}$. Without loss of generality, we may assume $(x, y) \in R$ and $(y, x) \notin R$. Then, it is obvious that $R \cap S^2 \subseteq S^1 \subseteq S^2 \cup R$ which means S^1 is between S^2 and R .

Case 2: \mathcal{X} contains incomplete or not antisymmetric relations. Then the type of elementary changes in the gapless simple subdomain \mathcal{S} consists of a single ordered pair. Hence we may assume, without loss of generality, that $S^1 \subseteq S^2$ and $S^1 \triangle S^2 = \{(y, x)\}$ for some $x, y \in A$. So $(y, x) \in S^2$ and $(y, x) \notin S^1$. Furthermore for any $(a, b) \neq (y, x)$, we have $(a, b) \in S^1$ if and only if $(a, b) \in S^2$.

Now, if $(y, x) \in R$, then $S^1 \cap R \subseteq S^2 \subseteq S^1 \cup R$ which means S^2 is between S^1 and R . If $(y, x) \notin R$, then $S^2 \cap R \subseteq S^1 \subseteq R \cup S^2$ which means S^1 is between S^2 and R . So, in all cases we find S^i is between R and S^j for some numbering i and j of $\{1, 2\}$. Lemma 6 now completes the proof. ■

Theorem 3 *Let \mathcal{X} be a domain of orders on A and \mathcal{S} a gapless simple subdomain such that all orders in \mathcal{X} are aligned towards \mathcal{S} . Let δ be a dissimilarity function on \mathcal{X} . Then*

1. δ satisfies betweenness if and only if there exists a weight function v such that $\delta = \kappa_v^*$.
2. δ satisfies betweenness and neutrality if and only if there exists a scalar $\alpha > 0$ such that $\delta = \kappa_\alpha^*$.
3. δ satisfies betweenness, neutrality, and normalization if and only if $\delta = \kappa^*$.

Proof. The “if-parts” are obvious and are left to the reader. Here we prove the “only-if-parts”.

(*Proof of 1*) Let R and R' be orders in \mathcal{X} . It is sufficient to prove $\delta(R, R') = \kappa_v^*(R, R')$. First we prove this for the case that R and R' are in \mathcal{S} . In view of Theorem 2 we may assume that on \mathcal{S} dissimilarity function δ equals κ_ω . In case \mathcal{X} consists of antisymmetric and complete relations only, by taking $v(a, b) = \frac{\omega(a, b) + \omega(b, a)}{2}$, we have that $\delta(R, R') = \kappa_\omega(R, R') = \kappa_v^*(R, R')$ for all orders in R and R' in \mathcal{X} . In case not all orders in \mathcal{X} are complete or not all orders in \mathcal{X} are anti-symmetric by Lemma 1 for all different alternatives a and b , we may find elementary changes of type $\{(a, b)\}$. Betweenness implies that the distance on such elementary changes is constant. So, we may take $v(a, b) = \delta(\hat{R}, \hat{R}')$, where \hat{R} and \hat{R}' form an elementary change of type $\{(a, b)\}$. Lemma 1 and betweenness imply that $\delta(R, R') = \kappa_v^*(R, R')$ for all R and R' in \mathcal{X} . So, the only if part of 1 holds for orders in \mathcal{S} .

Next we show the general case. Because all orders in \mathcal{X} are aligned towards \mathcal{S} we may assume further that R' is between R and some order S in \mathcal{S} . In view of Lemma 6 it is therefore sufficient to prove that $\delta(R, S) = \kappa_v^*(R, S)$ for all R in \mathcal{X} and all S in \mathcal{S} . By induction to $k \geq 1$ it is sufficient to prove that for all subsets X of A , with $\#X \leq k$, for all R in \mathcal{X} and all S in \mathcal{S}

$$\delta(R_X \gg L_{A \setminus X}, S_X \gg L_{A \setminus X}) = \kappa_v^*(R_X \gg L_{A \setminus X}, S_X \gg L_{A \setminus X}).$$

Basis $k \leq 2$. As in that case $R_X \gg L_{A \setminus X}$ and $S_X \gg L_{A \setminus X}$ are in \mathcal{S} it follows by the start of the proof that $\delta(R_X \gg L_{A \setminus X}, S_X \gg L_{A \setminus X}) = \kappa_v^*(R_X \gg L_{A \setminus X}, S_X \gg L_{A \setminus X})$.

Induction step Consider subset X of A , with $\#X = k+1 \geq 3$, $R \in \mathcal{X}$ and $S \in \mathcal{S}$. It is sufficient to prove $\delta(R_X \gg L_{A \setminus X}, S_X \gg L_{A \setminus X}) = \kappa_v^*(R_X \gg L_{A \setminus X}, S_X \gg L_{A \setminus X})$ under the assumption of the following induction hypothesis. For all subsets \hat{X} of A , with $\#\hat{X} \leq k$, for all \hat{R} in \mathcal{X} and all \hat{S} in \mathcal{S} we have $\delta(\hat{R}_{\hat{X}} \gg \hat{L}_{A \setminus \hat{X}}, \hat{S}_{\hat{X}} \gg \hat{L}_{A \setminus \hat{X}}) = \kappa_v^*(\hat{R}_{\hat{X}} \gg \hat{L}_{A \setminus \hat{X}}, \hat{S}_{\hat{X}} \gg \hat{L}_{A \setminus \hat{X}})$. For alternatives x in X let $R^x = R_{X \setminus \{x\}} \gg x \gg L_{A \setminus X}$.

Claim 1 $\delta(R^x, R_X \gg L_{A \setminus X}) = \kappa_v^*(R^x, R_X \gg L_{A \setminus X})$.

Proof of Claim 1 As $\#X \geq 3$ we may find at least three different alternatives in X , say a, b and c . It is sufficient to prove $\delta(R^a, R_X \gg L_{A \setminus X}) = \kappa_v^*(R^a, R_X \gg L_{A \setminus X})$.

First we prove that $\delta(R^a, R^b) = \kappa_v^*(R^a, R^b)$. As all orders in \mathcal{X} are aligned towards \mathcal{S} we may without loss of generality assume that R^b is between R^a and some order say S' in \mathcal{S} . Obviously $S'_{X \setminus \{a, b\}} \gg a \gg b \gg L_{A \setminus X}$ is in \mathcal{S} and between R^b and S' . So, we may assume that $S' = S'_{X \setminus \{a, b\}} \gg a \gg b \gg L_{A \setminus X}$. Also observe that $R_X \gg L_{A \setminus X}$ is between R^a and R^b . By these betweenness relations we have

$$\delta(R^a, R_X \gg L_{A \setminus X}) + \delta(R_X \gg L_{A \setminus X}, R^b) = \delta(R^a, R^b) \text{ and}$$

$$\delta(R^a, R^b) + \delta(R^b, S') = \delta(R^a, S').$$

Taking $S^b = S'$ and $S^a = S'_{X \setminus \{a, b\}} \gg b \gg a \gg L_{A \setminus X}$ the induction hypothesis implies

$$\begin{aligned} \delta(R^a, S^a) &= \kappa_v^*(R^a, S^a) \text{ and} \\ \delta(R^b, S^b) &= \kappa_v^*(R^b, S^b). \end{aligned}$$

As $\delta(R^a, S^a) = \kappa_v^*(R^a, S^a)$ and $S^b = S'$ and S^a are in \mathcal{S} it follows by Lemma 7 that $\delta(R^a, S') = \kappa_v^*(R^a, S')$. As R^b is between R^a and order $S^b = S'$, and $\delta(R^a, S') = \kappa_v^*(R^a, S')$ and $\delta(R^b, S^b) = \kappa_v^*(R^b, S^b)$ Lemma 6 yields $\delta(R^a, R^b) = \kappa_v^*(R^a, R^b)$.

Similarly we have $\delta(R^a, R^c) = \kappa_v^*(R^a, R^c)$ and $\delta(R^b, R^c) = \kappa_v^*(R^b, R^c)$.

To prove the claim it is sufficient to prove $\delta(R^a, R_X \gg L_{A \setminus X}) = \kappa_v^*(R^a, R_X \gg L_{A \setminus X})$. Recall that $R_X \gg L_{A \setminus X}$ is between R^a and R^b . So, we have

$$\begin{aligned} \delta(R^a, R_X \gg L_{A \setminus X}) + \delta(R_X \gg L_{A \setminus X}, R^b) &= \delta(R^a, R^b) = \\ &= \kappa_v^*(R^a, R^b) = \kappa_v^*(R^a, R_X \gg L_{A \setminus X}) + \kappa_v^*(R_X \gg L_{A \setminus X}, R^b) \end{aligned} \quad (7)$$

$$\begin{aligned} \delta(R^a, R_X \gg L_{A \setminus X}) + \delta(R_X \gg L_{A \setminus X}, R^c) &= \delta(R^a, R^c) = \\ &= \kappa_v^*(R^a, R^c) = \kappa_v^*(R^a, R_X \gg L_{A \setminus X}) + \kappa_v^*(R_X \gg L_{A \setminus X}, R^c). \end{aligned} \quad (8)$$

$$\begin{aligned} \delta(R^b, R_X \gg L_{A \setminus X}) + \delta(R_X \gg L_{A \setminus X}, R^c) &= \delta(R^b, R^c) = \\ &= \kappa_v^*(R^b, R^c) = \kappa_v^*(R^b, R_X \gg L_{A \setminus X}) + \kappa_v^*(R_X \gg L_{A \setminus X}, R^c). \end{aligned} \quad (9)$$

Adding Equations 7 and 8, and simplifying with Equation 9, while applying symmetry of dissimilarity δ to $\delta(R_X \gg L_{A \setminus X}, R^b)$, yields $2 \cdot \delta(R^a, R_X \gg L_{A \setminus X}) = 2 \cdot \kappa_v^*(R^a, R_X \gg L_{A \setminus X})$. So, $\delta(R^a, R_X \gg L_{A \setminus X}) = \kappa_v^*(R^a, R_X \gg L_{A \setminus X})$. This ends the proof of claim 1.

Consider $S^x = S_{X \setminus \{x\}} \gg x \gg L_{A \setminus X}$ for some x in X . Order S^x is in \mathcal{S} and further R^x is between $R_X \gg L_{A \setminus X}$ and S^x . So,

$$\delta(R_X \gg L_{A \setminus X}, S^x) = \delta(R_X \gg L_{A \setminus X}, R^x) + \delta(R^x, S^x).$$

The claim and the induction hypothesis yield $\delta(R_X \gg L_{A \setminus X}, S^x) = \kappa_v^*(R_X \gg L_{A \setminus X}, R^x) + \kappa_v^*(R^x, S^x) = \kappa_v^*(R_X \gg L_{A \setminus X}, S^x)$. Lemma 7 now implies that $\delta(R_X \gg L_{A \setminus X}, S_X \gg L_{A \setminus X}) = \kappa_v^*(R_X \gg L_{A \setminus X}, S_X \gg L_{A \setminus X})$.

(Proof of 2) Follows identical to the proof of Theorem 2.2

(Proof of 3) Follows identical to the proof of Theorem 2.3. ■

Theorem 3 extends the characterization of Kemeny distance as can be found in Kemeny (1959) in several ways. Obviously it extends this characterization to larger domains of orders. Also the characterization above only uses a fixed set of alternatives where orders on all non-empty subsets of alternatives are needed in Kemeny's characterization. Another logical extension is implied by the redundancy of reducibility.

The following example shows that the assumption on all orders in \mathcal{X} being aligned towards a gapless simple subdomain \mathcal{S} is essential. That is, dropping it may lead to domains on which betweenness does not yield weighted Kemeny distance. Below we provide an example of a domain which is not aligned and hence the characterization does not work.

Example 6 Let A consist of four alternatives. Take domain \mathcal{X} consisting of two disjoint subsets \mathcal{C} and \mathcal{A} where \mathcal{C} is a set of cyclic relations on A , and \mathcal{A} is a set of acyclic relations on A , specified as follows:

$\mathcal{C} = \{A \times A \setminus \{(a, b), (b, c), (c, d), (d, a)\} : \text{for some naming } a, b, c \text{ and } d \text{ of all the alternatives in } A\}$ and $\mathcal{A} = \{L \cup \{(a, b)\} : L \in \mathcal{L} \text{ and } a, b \in A\}$.

Note that \mathcal{C} consists of all cycles on four alternatives including two additional disjoint indifferences¹⁴ and \mathcal{A} consists of all relations obtained from a linear order by adding one more ordered

¹⁴Consider for instance the cycle $C \in \mathcal{C}$ on $\{a, b, c, d\}$ such that $(d, c), (c, b), (b, a), (a, d) \in C$. C also includes the ordered pairs, $(d, b), (b, d), (c, a)$, and (a, c) , which constitutes the two disjoint indifferences.

pair¹⁵. Note that this additional ordered pair, e.g., (x, y) may lead to either i) the same order¹⁶, or ii) a weak order¹⁷, or iii) an intransitive relation¹⁸. So, relations in \mathcal{A} has at most one “real” indifference. Consider an arbitrary relation C in \mathcal{C} and an arbitrary relation R in \mathcal{A} . Then the symmetric set difference between C and R contains at least two ordered pairs of alternatives. One to break the strict four-cycle in C and an other to tie-break one of the two indifferences in C .

By definition, \mathcal{X} consists of reflexive relations only and further it is closed under permutation. Applying linearization to elements in \mathcal{C} yields elements in \mathcal{A} , where applying linearization to elements in \mathcal{A} yields elements in \mathcal{A} . So, \mathcal{X} satisfies the conditions of Definition 1 and is therefore a proper domain. Next, we show that there is no gapless simple subdomain \mathcal{S} of \mathcal{X} towards which \mathcal{X} is aligned. So we show that \mathcal{X} is “not” an aligned domain.

We first show that a gapless simple subset of \mathcal{X} is in fact a subset of \mathcal{A} . Let \mathcal{S} be a gapless simple subdomain of \mathcal{X} . Note that $\mathcal{L}^\cup \subseteq \mathcal{X}$. By this it follows that $\mathcal{L}^\cup \subseteq \mathcal{S} \subseteq \mathcal{X}$ and that the types of elementary changes in \mathcal{S} are singletons. Since the symmetric set difference between orders in \mathcal{C} with those in \mathcal{A} contain at least two ordered pairs of alternatives we have that \mathcal{S} cannot contain orders which are in \mathcal{C} . So, $\mathcal{L}^\cup \subseteq \mathcal{S} \subseteq \mathcal{A} \subseteq \mathcal{X}$.

Now let two cycles in \mathcal{C} oriented in opposite direction, e.g.:

$$C^1 = A \times A \setminus \{(b, a), (a, d), (d, c), (c, b)\}$$

$$C^2 = A \times A \setminus \{(a, b), (b, c), (c, d), (d, a)\}.$$

Consider the intersection of these; it contains $\{(a, c), (c, a), (b, d), (d, b)\}$. That is, the indifference between a and c as well as the indifference between b and d . Because none of the other relations in \mathcal{X} contain both these two indifferences simultaneously, it follows that none of the other relations in \mathcal{X} is between C^1 and C^2 . Hence, C^1 and C^2 form an elementary change.

To prove that \mathcal{X} is not aligned towards any \mathcal{S} , take order R in \mathcal{A} . It is sufficient to show that i) C^2 is not between R and C^1 , and ii) C^1 is not between R and C^2 . We only prove (i) as (ii) follows similarly. Note that by definition, all orders in \mathcal{A} have i) a unique best alternative or ii) a unique worst.

- If $R \in \mathcal{S} \subseteq \mathcal{A}$ has a unique “best”, say a , then at R and C^1 alternative a is strictly ordered above alternative b . Thus, the intersection of C^1 and R is not contained in C^2 .
- If $R \in \mathcal{A}$ has a unique “worst”, say a , then at R and C^1 alternative d is strictly ordered above alternative a . Thus, the intersection of C^1 and R is not contained in C^2 .

This implies in either case, i.e., R having a unique best or a unique worst at R , that C^2 is not between C^1 and R . So, elementary change $\{C^1, C^2\}$ is not aligned towards any $R \in \mathcal{A}$. Therefore betweenness has no inference power on the value of such elementary changes.

Finally we introduce a dissimilarity function on this domain. So, as to not disturbing triangle inequality, we may define δ for elementary changes such as C^1 and C^2 just an epsilon larger or smaller than Kemeny distance and keep Kemeny distance for all the other elementary changes. This yields a dissimilarity function which is neutral, normalized and satisfies betweenness. It is obviously not equal to Kemeny distance.

¹⁵Consider for instance $R \in \mathcal{A}$ which is the union of the linear order $L = a \gg b \gg c \gg d$ and an ordered pair (c, a) .

¹⁶In case the pair is already in that linear order

¹⁷In case y is consecutively and strictly ordered above x at that linear order

¹⁸In case y is ordered strictly above x but not consecutively

We end this section by demonstrating a weighted Kemeny distance as an example below:

Example 7 Let $A = \{a, b, c\}$. Consider the domain of weak orders in Figure 7 and a weight distribution v as follows: $v(b, a) = \alpha_1 = 3.5, v(a, b) = \alpha_2 = 2, v(c, b) = \beta_1 = 2.5, v(b, c) = \beta_2 = 3, v(c, a) = \gamma_1 = 4.5, v(a, c) = \gamma_2 = 4$. Then the distance κ_v between each elementary change is depicted below.

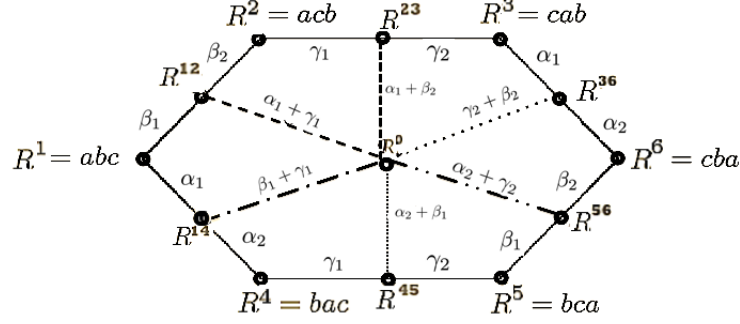


Figure 4: A graph for weak orders when $m = 3$. Given the v weights in Example 7, betweenness necessitates that $\kappa^v(R^1, R^0) = \beta_1 + \gamma_1 + \alpha_1 = 10.5$, $\kappa^v(R^1, R^{14}) = \alpha_1 = 3.5$, $\kappa^v(R^1, R^{12}) = \beta_1 = 2.5$, $\kappa^v(R^{14}, R^0) = \beta_1 + \gamma_1 = 7$, $\kappa^v(R^{12}, R^0) = \alpha_1 + \gamma_1 = 8$, $\kappa^v(R^1, R^6) = \alpha_1 + \beta_1 + \gamma_1 + \alpha_2 + \beta_2 + \gamma_2 = 19.5$.

6 Discussion

We have generalized well established measures on orders by first providing an extensive framework on orders and thereafter characterizing various classes of dissimilarity (distance) functions on these orders. Our results on how to compare orders can be applied to various scenarios as also indicated in the introduction, e.g., strict rankings, weak orders, tournaments etc. In the first case a scenario where positions of disagreements among orders are weighted is thoroughly analyzed and in the second case a pairwise weighted approach is discussed. In this latter case it appears that betweenness structures the dissimilarity function considerably. Neutrality is sufficient to single out scaled Kemeny distances from weighted ones, but much less demanding conditions can establish the same. Hence, essentially we have shown that under betweenness the classes of dissimilarity functions can be narrowed down to the Kemeny distance by imposing some mild axioms on these measures.

Next we discuss the logical independence of the characterizing conditions and thereafter present the connection between basic metric conditions and the properties we employ in this paper. Finally we dwell on possible future work.

6.1 Logical independence of the characterizing conditions

The discussion on the logical independence of the various characterizing conditions in Theorems 2 and 1 (and in Corollary 1) needs two rather technical dissimilarity functions. These dissimilarity

functions are:

$$\kappa_\varepsilon(R, R') = \begin{cases} \kappa(R, R') - \varepsilon, & \text{if } R' = -R \\ \kappa(R, R'), & \text{otherwise} \end{cases}$$

$$\delta_?(R, R') = \max \left\{ \sum_{i=1}^l \tau(E\{R^i, R^{i+1}\}) : [R^1, \dots, R^{l+1}] \text{ is a path of elementary changes from } R \text{ to } R' \text{ with } l \text{ strictly positive.} \right\}$$

It is elementary to check which of the conditions introduced in the foregoing sections are satisfied by these dissimilarity functions. The following table reflects the outcome of this. A “+” sign indicates that the function satisfies the condition and a “-” sign that it does not. We list the conditions as 1) TI: triangular inequality, 2) BET: betweenness, 3) DEC: decomposability, 4) IND: independence, 5) NEU: neutrality, 6) NOR: normalization, 7) PAI: independence of pair addition

	κ_ε	$\delta_?$
TI	+	-
BET	-	-
DEC	-	+
IND	+	-
NEU	+	+
NOR	+	+
PAI	+	-

The logical independencies of the conditions in Theorems 2, 1 and Corollary 1 are now to be found in the following table. Here a “+” sign indicates that all dissimilarity functions in the related class satisfy this condition and a “0” sign indicates that some of these do and some of these don’t. In case a condition is one of the characterizing conditions of the class at hand, this is indicated by a superscript referring to the Theorem (or Corollary) part describing this characterization by **T** (or respectively **C**). In some case two superscripts may be found meaning that the condition is a characterizing condition in two theorems/corollaries. In the last row of this table the stacks of functions summarize the logical independencies (denoted as log. ind.). That is the i^{th} element of the stack (top to bottom) is an example of a dissimilarity function satisfying all but the i^{th} characterizing condition.

	κ_ω	κ^τ	κ_α	κ
TI	+	$+\mathbf{T1}$	+	+
BET	$+\mathbf{T2.1}$	0	$+\mathbf{T2.2}$	$+\mathbf{T2.3}$
DEC	$+\mathbf{C(A\&B)1.1}$	$+\mathbf{T1}$	$+\mathbf{C(A\&B)1.2}$	$+\mathbf{C(A\&B)1.3}$
IND&PAI	$+\mathbf{C(A\&B)1.1}$	0	$+\mathbf{C(A\&B)1.2}$	$+\mathbf{C(A\&B)1.3}$
NEU	0	$+\mathbf{T1}$	$+\mathbf{T2.2;C(A\&B)1.2}$	$+\mathbf{T2.3;C(A\&B)1.3}$
NOR	0	0	0	$+\mathbf{T2.3;C(A\&B)1.3}$
	$\mathbf{T2.1}$ $\mathbf{C(A\&B)1.1}$	$\mathbf{T1}$	$\mathbf{T2.2}$ $\mathbf{C(A\&B)1.2}$	$\mathbf{T2.3}$ $\mathbf{C(A\&B)1.3}$
log.	\cdot	$\delta?$	\cdot	\cdot
ind.	κ^τ	\cdot	κ^τ	κ^τ
	\cdot	κ_ε	\cdot	κ_ε
	\cdot	κ^τ	\cdot	κ^τ
	\cdot	κ_ω	κ_ω	κ_ω
	\cdot	\cdot	\cdot	κ_α

All in all we may conclude that the characterizing conditions in the various parts of Theorems 1, 2, and Corollary 1 are all logically independent.

6.2 The basic conditions for dissimilarity functions

In this subsection we study the logical relations between the sub-conditions which define a dissimilarity function and the conditions of betweenness, decomposability, independence, neutrality, normalization and independence of pair addition used in the characterization of the Kemeny and Kendall distance. Therefore, we model a pseudo-dissimilarity function between orders by an arbitrary function instead of one satisfying non-negativity and symmetry. Similar to Bogart (1973), we consider *pseudo-dissimilarity* functions, say λ , defined for tuples of orders in \mathcal{X} . For orders R and R' in \mathcal{X} let $\lambda(R, R')$ be a real number indicating the *pseudo-dissimilarity* between R and R' . Note that the similar conditions imposed on distance functions apply to λ . To avoid trivial reformulation, however, we refer to these by the same names and numbers as we did so far for dissimilarity functions.

It is clear from the previous subsection that the characterizing conditions triangle-inequality, betweenness, independence, decomposability, neutrality and normalization in the Theorems 2, 1 and Corollary 1 are all logically independent of the two basic conditions non-negativity (denoted as NN) and symmetry (SYM) of the dissimilarity function. By the following table we will show the reverse, with one exception. It is elementary to prove that symmetry and triangle inequality imply non-negativity¹⁹. That is these two dissimilarity function defining conditions are logically independent of all the other seven mentioned conditions. The following pseudo-dissimilarity functions are needed to show this independence.

Let \mathcal{X} be a set of orders on A with dichotomous elementary change structure. Let ν be a function on \mathcal{X} defined by

$$\begin{aligned} \nu(R, R') = & 3 \cdot \#\{(x, y) \in R : (y, x) \notin R \text{ and } (x, y) \notin R' \text{ and } (y, x) \in R'\} + \\ & 2 \cdot \#\{(x, y) \in R : (y, x) \in R \text{ and } (x, y) \notin R' \text{ and } (y, x) \in R'\} + \\ & 1 \cdot \#\{(x, y) \in R : (y, x) \notin R \text{ and } (x, y) \in R' \text{ and } (y, x) \in R'\}. \end{aligned}$$

¹⁹Triangle inequality implies $\lambda(a, a) + \lambda(a, a) \geq \lambda(a, a)$. So, $\lambda(a, a) \geq 0$. Triangle inequality implies $\lambda(a, b) + \lambda(b, a) \geq \lambda(a, a)$. So, $\lambda(a, b) + \lambda(b, a) \geq 0$ and symmetry now implies $2 \cdot \lambda(a, b) \geq 0$. So, $\lambda(a, b) \geq 0$.

So, like at Kemeny distance ν takes into account the pairs of alternatives which have to be changed to obtain order R' from R . Where each change from a strict order of alternative x above alternative y at R to a strict reversed order of y above x at R' contributes 3 to the outcome $\nu(R, R')$. And, each change from an indifference between x and y at R to a strict order of y above x at R' contributes 2 and each change from a strict order of x above y at R to an indifference between these two at R' contributes 1 to $\nu(R, R')$. Note that ν is not symmetric in its arguments. Therefore ν does not satisfy symmetry of the metric condition.

	ν	$-\kappa$
NN	+	−
SYM	−	+
TI	+	−
BET	+	+
DEC	+	+
IND	+	+
NEU	+	+
NOR	+	+
PAI	+	+

6.3 Future work

There are three natural routes to further investigate dissimilarities between orders. One is on incorporating infinite sets of alternatives. This obviously would create the absence of elementary changes. So, the proofs of our main results cannot be updated straightforwardly in this direction. Of course Lemma 3 would still be applicable and may be of help to deal with independence. Another route would more or less go in perpendicular direction and study on finite sets of alternatives and subsets of orders such as single peaked orders or single caved ones. Now the definition of a domain has to be altered. For instance it cannot longer be closed under permutation and conditions like neutrality can no longer be imposed. Finally the last route would lead to a thorough search on all kind of relaxations and alterations of conditions like neutrality, independence (of pair addition) and would therewith stick closest to the findings presented above.

References

- Bogart, K. (1973). Preference structures I: Distances between transitive preference relations. *Journal of Mathematical Sociology* 3(1), 49–67.
- Can, B. (2012). *Essays in Microeconomic Theory*. Ph. D. thesis, Maastricht University.
- Can, B. (2014). Weighted distances between preferences. *Journal of Mathematical Economics* 51, 109–115.
- Can, B. and T. Storcken (2013). A re-characterization of the Kemeny distance. Technical report, Maastricht: GSBE, Graduate School of Business and Economics.
- Cramer, G. (1750). *Introduction à l'analyse des lignes courbes algébriques*. Genève: Frères Cramer et C. Philibert.

- Farnoud, F. and O. Milenkovic (2012). Sorting of permutations by cost-constrained transpositions. *Information Theory, IEEE Transactions on* 58(1), 3–23.
- Farnoud, F., B. Touri, and O. Milenkovic (2012). Novel distance measures for vote aggregation. *arXiv preprint arXiv:1203.6371*.
- Hamming, R. (1950). Error detecting and error correcting codes. *Bell System Technical Journal* 29(2), 147–160.
- Kemeny, J. (1959). Mathematics without numbers. *Daedalus* 88(4), 577–591.
- Kemeny, J. and J. Snell (1962). *Mathematical models in the social sciences*. Boston: Ginn.
- Kendall, M. (1938). A new measure of rank correlation. *Biometrika* 30(1/2), 81–93.
- Knuth, D. (1998). *The art of computer programming: sorting and searching*, Volume 3. Reading, MA: Addison Wesley Longman Publishing Co., Inc.
- Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica, Journal of the Econometric Society*, 178–191.
- Storcken, T. and H. de Swart (1992). Towards an axiomatization of orderings. *Order* 9(4), 367–396.